MS&E 233: Networked Markets - Project
Exchanges in Online Advertising

Misrab Faizullah-Khan
Stanford University

June 2012

1 Introduction

We are now well into a multi-decade long transition from printed to online readership. A myriad of publications are now available online, sometimes even rendering the printed version obsolete. This has allowed for a dynamic online advertising market to flourish, where users can now be served with advertisements in real-time.

Ad serving technology that stores advertisers' content and delivers them to webpages when required has developed in tandem with this. [1] The scale of this market can be highlighted by glancing at the number of viewers catered to by the leading ad servers in 2008, in what is a relatively dense oligopoly: [2]

Vendor	Ad Viewers (millions)
Google	1,118
DoubleClick (Google)	1,079
Yahoo!	362
MSN (Microsoft)	309
AOL	156
Adbrite	73

In this paper, we will examine one of the processes through which a user is served a certain advertisement when visiting a webpage. We will notice the importance of online advertising exchanges, and therefore take the perspective of such an exchange trying to maximize revenue. We will be able to define this question much more specifically once we have introduced the model in the next section.

2 Modeling Assumptions

2.1 Advertising Exchange Model

We will base our study off the model of Ad Exchanges presented by S. Muthukrishnan of Google Inc. [3] In this model, the following steps are undergone to present an ad in a single slot:

1. User with information u visits webpage w of publisher $p(w)$.
2. Publisher $p(w)$ sends (u, w, p) to the exchange E, where p is the minimum price it will accept for the given slot.
3. The exchange E then contacts ad networks $a_1, \ldots, a_m$ that represent advertisers, with a modulated $(E(u), E(w), E(p))$ determined by E.
4. Each ad network i then returns a bid on behalf of its advertiser customers, and the desired ad respectively as (b_i, d_i).
5. E runs an auction on these, of which the winner is (b_i^*, d_i^*), who pays c_i^*, where $p \leq c_i^* \leq b_i^*$.
6. Finally, the publisher $p(w)$ received the tuple (c_i^*, d_i^*) and serves the designated ad to user u.

Note that this model will be slightly extended in the course of our analysis, and we will highlight these changes. What we will seek to answer is how the exchange should determine $E(u), E(w)$ and $E(p)$ given u, w and p in order to maximize revenue. We will of course have to contextualize this within relevant assumptions, which we make as follows.

2.2 Engineering Assumptions

1. How does E make profit? We assume it does so solely through a commission αc_i^*, $\alpha \in [0, 1]$.
2. E selects a winning bid from $a_i, \ldots, a_m$ through a second-price auction. The motivation for this is, as studied, it leads to truthful bidding by ad exchanges on behalf of their customers. We will not delve into the mechanism between $a_1, \ldots a_m$ and their customers, leading to our next assumption. [4]
3. Ad networks mirror their customers: they share identical valuations, and identical preferences for user attributes. In fact, we can consider an ad network to actually be its own unique customer for simplicity.

2.3 Behavioral Assumptions

1. As stated, E wishes to maximize profit. We will also assume this takes into account both the short run and long run; the exchange will not pursue a de facto myopic or "greedy" policy, in the algorithmic sense of the word, at the cost of future transactions.
2. $p(w)$ seeks to maximize profit as above. This entails a high per transaction profit quantitatively, but also qualitatively a sustainable relationship with the exchange over multiple time periods. The former includes expectations, which we will explore further.
3. $a_1, \ldots, a_m$ act as studied in a second-price auction, with private values and a goal to maximize expected payoff.
4. For completeness, we will also assume that user u prefers certain ads over others. This may or may not translate into a higher click through rate in the short run, but certainly not a lower one. Furthermore, it has qualitative benefits to $p(w)$ in the long run.

In the next sections, we will explore what kind of information the exchange E receives in (u, w, p), how $a_1, \ldots, a_m$ would react to this information, and consequently how E might want to modulate it. More subtle assumptions within arguments will be clearly stated.

3 User Information

Let us first determine what unshaded information $p(w)$ might have regarding u. Call this u', from which $p(w)$ may choose to omit certain attributes when it passes u to E.

3.1 Passive Information

This denotes what information $p(w)$ readily has on the user upon a visit, before the user takes any actions on the site w. Typical elements include:

1. Time of visit;
2. IP address and general location (subject to proxies);
3. Revealed preference given visit non-random: language, interest in content;
4. Based on IP address: passive information from past visits (i.e. time patterns...)
5. If logged in through a social network (i.e. Facebook...), possibly social graph information. This has provided much richer passive information on users in recent years.

3.2 Active Information and Revealed Preference

This includes the activity of the user that can be detected on the website. It is clearly very site-specific, and can range from commercial transactions, to navigation or social interaction.

It also includes past active information retrieved from the user and stored.

3.3 Consequences on Exchange's Decision

Remember that all of this information can have either a negative of positive effect on bids. There are several key points to note then. First of all, there is one scenario where $p(w)$ has some knowledge of how E modulates u, and another with $p(w)$ doesn't. Readings suggest that in the case where $p(w)$ and E are indeed separate entities, the amount of information the former has on the latter is very limited. [5]

If we make the strong assumption then that $p(w)$ has no information on E's mechanism, we can consider $p(w)$'s incentives more clearly. We now show that $p(w)$ will choose to either reveal more information about u consistently, or less information consistently. This is because if the amount of information is a high H when $p(w)$ believes it would lead to higher bidder valuations (or likely to be so with probability $p > 1/2$), and a low L otherwise, we have:

$$P(H|good) = p > 1/2, P(L|good) = (1 - p) < 1/2$$

Baye's rule then gives the ad networks a high $P(good|H)$, and likewise for the "bad" user case. This is essentially then a signaling game as studied in class, which dampens much of the efficacy of $p(w)$ concealing information methodically. [6]

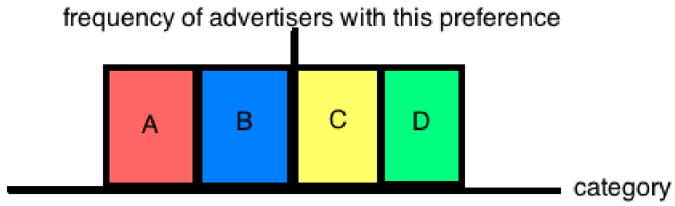
Now consider the case in which $p(w)$ is deciding whether to shade or reveal u (partially), albeit consistently. Consider illustrations of the positive and negative effects of the attributes of u mentioned in the previous sections on an advertiser's valuation (given ad networks mirror advertisers, as per our engineering assumptions):

Attribute	Positive	Negative
time of visit	suitable for a given ad e.g. meal time for take-out	inappropriate time
location	close the given advertiser, or suitable geographical market	different country from advertiser
past passive information	high value repeat customer	random visiting or inappropriate time consistently
social graph	correct market segment	incorrect segment or non-influential node
active information	past click-through rates and ads	transacted with competitor or incompatible item basket

At this point, it seems somewhat arbitrary whether more information on the user u would lead to a higher or lower bid. We will lean towards the former based on two approaches: one empirical, the other theoretical.

In the first case, the publisher $p(w)$ may have access to historical data that suggests a certain conclusion. Barring that, we can look at more general data. Although pure statistics on advertiser valuations given user information seem scarce, one paper on the DoubleClick Ad exchange, acquired by Google, suggests that increased information on u resulted in higher revenues for the ad exchange.[7] Based on our assumptions, if the exchange E's unit transaction revenue αc_i^* is higher, then $(1-\alpha)c_i^*$ is proportionally higher and $p(w)$ too has a higher unit revenue. If the revenue is increased by transaction volume instead, then publishers on aggregate benefit, although the distribution of revenue among publishers still remains ambiguous.

Given the limited data, we can take a theoretical approach to gain some insight. The key idea is that user attributes might not be a binary positive or negative, but rather simply indicative or which "type" of consumer they are. For illustration, assume there are four categories of users (market segments): A, B, C and D. The deciding factor will be the distribution with which advertisers prefer consumers of each category. If we assume a normal distribution, with say more advertisers preferring B and C type consumers, then more information on u should be revealed if there are more of these consumers on average. If we assume a uniform distribution of advertiser preferences across categories as illustrated below, then on average a user less appealing to one advertiser may seem more appealing to another:



In this case, we would expect a higher valuation $v_i' > v_i$ for some advertiser. If $v_i' > v_j, j = 1, \ldots, m$, then given truthful bidding revenue for both $p(w)$ and E would increase.

Under this conclusion, $p(w)$ would pass a complete u to E, and $E(u) = u$.

4 Webpage Information

Let's now study how E might want to modulate information about the publisher's webpage, w. We start by considering what information might be contained therein, and how this might affect bids and in consequence $p(w)$'s and E's behaviors.

Since $p(w)$ is conveying the information to E, we allow $w \subset w'$ where more information could be willingly provided by $p(w)$ about its webpage, and w denotes non-concealable information to E (i.e. what can be extracted directly from the webpage).

4.1 Revealed and Revealable Information

We'll include potential attributes taken from both w and $y = w' \backslash \{w\}$ in this section, and examples on how they could have positive or negative effects on potential bids:

Attribute	Positive	Negative
global content of page (w)	site draws users suitable to certain ads (revealed preference)	users not suitable to ads
local content near ad slot (w)	article on benefits of coffee for coffee advertiser	article on carcinogenic properties of coffee
listed information on $p(w)$ (w or y)	(limited)	bad press, to which ads should not be associated
site statistics and click-through rates (y)	positive correlation	negative correlation
site-collected statistics on users (y)	user segmentation	unsuitable user segments

Note that above examples extend Muthukrishnan's model, where it is understood w merely represents the webpage address. Here we are allowing for much richer information to be passed, which is already often the case with online services. [8]

4.2 Consequences on Exchange's Decision

The arguments on the previous section on u would suggest that more information on w would signal a higher potential value of viewers to ad networks. However, the analysis here differs in one key point: some statistics are perfectly linear, and not categorical as with users. What we mean by this is there is more at play than simple horizontal market segmentation: statistics on lower click-through rates and the like on a given webpage are outright negatively correlated to bids.

Therefore, although a certain $p(w)$ may wish to pass good statistics in y on to E, it will not do so with bad statistics. E must however adopt a consistent modulation due to the signaling effect on ad networks described earlier.

Following this logic, it is likely that the revenue-maximizing scenario for E is to consistently modulate w' such that $E(w') = w$. We will see that given the second-price auction, modulation is very simple here, although there are some interesting points in how $p(w)$ might behave.

5 Minimum Price

Finally, we look at how $p(w)$'s minimum acceptable price is conveyed to E and subsequently to $a_1, \ldots, a_m$.

5.1 Publisher's Incentives

First, consider the case without an exchange where $p(w)$ is selling the item. This is also the case where $p(w)$ does not know how E modulates p. Then $p(w)$ first declares p, the $a_1, \ldots, a_m$ independently state their bids with this knowledge, and the highest c_i^* wins. Suppose $p(w)$ has an opportunity cost $c \geq 0$ for the advertising slot. If $p(w)$ sets a minimum price below c, and some bid $c_i^* < c$ is returned, then $p(w)$ has a negative payoff of $c_i^* - c < 0$. Likewise, the payoff is 0 if $c = c_i^*$, and positive if $c < c_i^*$.

Let us consider the expectations involved in this last case where $p(w)$ declares $p > c$. The higher p, the higher the payoff if not all bids are null. The flip side is that without any a priori information on the independent private values of $a_1, \ldots, a_m$, a higher p increases the risk of all bids being null, resulting in a $-c$ payoff. The outcome will depend on $p(w)$'s assumption on the underlying distribution of bids (possibly supported by data), and its risk appetite.

5.2 Exchange's Decision

The above will have little bearing on E's modulation of p however. Recall that the very purpose of the second-price auction is to incentivize ad networks to bid truthfully. The alternative would be to go for gold: sell the item in hope of receiving the highest value of all $a_1, \ldots, a_m$.

Barring this risky method, the second price auction allows a winning bid of the second highest valuation. If this more consistent method is indeed selected, E will set $E(p) = \varnothing$.

6 Conclusion and Afterthoughts

Within the limited scope of this paper, it is likely that E will modulate its input from $p(w)$ as $E(u, w', p) = (u, w, \varnothing)$ with the symbols as defined.

With regards to each of these choices, more research could be pursued on issues including alternative revenue models for E, long-run game theoretic interactions between the parties at work and between ad networks and their customers, and alternative bidding mechanisms.

The topic of advertising exchanges is rich and complex, but hopefully this paper sheds some light on one of the design issues involved in this fascinating process.

Notes

[1]Emediate - What is Ad Serving? Available on the World Wide Web at: http://dataninja.wordpress.com/2006/06/06/endnotes-in-latex/

[2]Browser Media - DoubleClick Deal. Available on the World Wide Web at: http://www.browsermedia.co.uk/2008/04/01/doubleclick-deal-means-google-controls-69-of-the-online-ad-market/

[3]AdX: A Model for Ad Exchanges. Google, Inc.

[4]Networks, Crowds, and Markets, David Easley and Jon Kleinberg, Chapter 9: Auctions.

[5]Monetize: Ad Network vs Mediator vs Exchange. Available on the World Wide Web at: http://blog.mobclix.com/2010/11/10/monetize-ad-network-vs-mediator-vs-exchange-and-what-it-means-for-you/

[6]Networks, Crowds, and Markets, David Easley and Jon Kleinberg, Chapter 22: Markets and Information.

[7]myThings and Google: Combining the power of real-time bidding, personalized retargeting and Google's DoubleClick Ad Exchange.

[8]Personalized ads, M.I.T. Press. Available on the World Wide Web at: http://www.internetretailer.com/2011/06/03/personalized-ads-might-carry-less-power-previously-thought